RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _10/511,796A_

Source: _Pu7/o_

Date Processed by STIC: _5/19/05_

ENTERED

PCT

```
  4 <110> APPLICANT: Sode, Koji
  6 <120> TITLE OF INVENTION: GLUCOSE DEHYDROGENASE BETA-SUBUNIT AND
  7       DNA ENCODING THE SAME
  9 <130> FILE REFERENCE: TOYA126.002APC
 11 <140> CURRENT APPLICATION NUMBER: 10/511,796A
 12 <141> CURRENT FILING DATE: 2004-10-19
 14 <150> PRIOR APPLICATION NUMBER: PCT/JP03/05375
 15 <151> PRIOR FILING DATE: 2003-04-25
 17 <150> PRIOR APPLICATION NUMBER: JP 2002-125353
 18 <151> PRIOR FILING DATE: 2002-04-26
 20 <160> NUMBER OF SEQ ID NOS: 19
 22 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 2467
 26 <212> TYPE: DNA
 27 <213> ORGANISM: Burkhorderia cepacia
 29 <220> FEATURE:
 30 <221> NAME/KEY: CDS
 31 <222> LOCATION: (258)...(761)
 33 <221> NAME/KEY: CDS
 34 <222> LOCATION: (764)...(2380)
 36 <221> CDS
 37 <222> LOCATION: (2386)...(2466)
 39 <400> 1
 40 aagctttctg tttgattgca cgcgattcta accgagcgtc tgtgaggcgg aacgcgacat 60
 41 gcttcgtgtc gcacacgtgt cgcgccgacg acacaaaaat gcagcgaaat ggctgatcgt 120
 42 tacgaatggc tgacacattg aatggactat aaaaccattg tccgttccgg aatgtgcgcg 180
 43 tacatttcag gtccgcgccg atttttgaga aatatcaagc gtggttttcc cgaatccggt 240
 44 gttcgagaga aggaaac atg cac aac gac aac act ccc cac tcg cgt cgc     290
 45                     Met His Asn Asp Asn Thr Pro His Ser Arg Arg
 46                      1               5                  10
 48 cac ggc gac gca gcc gca tca ggc atc acg cgg cgt caa tgg ttg caa   338
 49 His Gly Asp Ala Ala Ala Ser Gly Ile Thr Arg Arg Gln Trp Leu Gln
 50          15                  20                  25
 52 ggc gcg ctg gcg ctg acc gca gcg ggc ctc acg ggt tcg ctg aca ttg   386
 53 Gly Ala Leu Ala Leu Thr Ala Ala Gly Leu Thr Gly Ser Leu Thr Leu
 54          30                  35                  40
 56 cgg gcg ctt gca gac aac ccc ggc act gcg ccg ctc gat acg ttc atg   434
 57 Arg Ala Leu Ala Asp Asn Pro Gly Thr Ala Pro Leu Asp Thr Phe Met
 58          45                  50                  55
 60 acg ctt tcc gaa tcg ctg acc ggc aag aaa ggg ctc agc cgc gtg atc   482
 61 Thr Leu Ser Glu Ser Leu Thr Gly Lys Lys Gly Leu Ser Arg Val Ile
 62 60                  65                  70                  75
```

```
64 ggc gag cgc ctg ctg cag gcg ctg cag aag ggc tcg ttc aag acg gcc   530
65 Gly Glu Arg Leu Leu Gln Ala Leu Gln Lys Gly Ser Phe Lys Thr Ala
66              80                  85                  90
68 gac agc ctg ccg cag ctc gcc ggc gcg ctc gcg tcc ggt tcg ctg acg   578
69 Asp Ser Leu Pro Gln Leu Ala Gly Ala Leu Ala Ser Gly Ser Leu Thr
70              95                 100                 105
72 cct gaa cag gaa tcg ctc gca ctg acg atc ctc gag gcc tgg tat ctc   626
73 Pro Glu Gln Glu Ser Leu Ala Leu Thr Ile Leu Glu Ala Trp Tyr Leu
74             110                 115                 120
76 ggc atc gtc gac aac gtc gtg att acg tac gag gaa gca tta atg ttc   674
77 Gly Ile Val Asp Asn Val Val Ile Thr Tyr Glu Glu Ala Leu Met Phe
78      125             .     130                 135
80 ggc gtc gtg tcc gat acg ctc gtg atc cgt tcg tat tgc ccc aac aaa   722
81 Gly Val Val Ser Asp Thr Leu Val Ile Arg Ser Tyr Cys Pro Asn Lys
82 140                 145                 150                 155
84 ccc ggc ttc tgg gcc gac aaa ccg atc gag agg caa gcc tg atg gcc    769
85 Pro Gly Phe Trp Ala Asp Lys Pro Ile Glu Arg Gln Ala   Met Ala
86             160                 165                 170
88 gat acc gat acg caa aag gcc gac gtc gtc gtc gtt gga tcg ggt gtc   817
89 Asp Thr Asp Thr Gln Lys Ala Asp Val Val Val Val Gly Ser Gly Val
90             175                 180                 185
92 gcg ggc gcg atc gtc gcg cat cag ctc gcg atg gcg ggc aag gcg gtg   865
93 Ala Gly Ala Ile Val Ala His Gln Leu Ala Met Ala Gly Lys Ala Val
94             190                 195                 200
96 atc ctg ctc gaa gcg ggc ccg cgc atg ccg cgc tgg gaa atc gtc gag   913
97 Ile Leu Leu Glu Ala Gly Pro Arg Met Pro Arg Trp Glu Ile Val Glu
98      205                 210                 215
100 cgc ttc cgc aat cag ccc gac aag atg gac ttc atg gcg ccg tac ccg  961
101 Arg Phe Arg Asn Gln Pro Asp Lys Met Asp Phe Met Ala Pro Tyr Pro
102      220                 225                 230
104 tcg agc ccc tgg gcg ccg cat ccc gag tac ggc ccg ccg aac gac tac 1009
105 Ser Ser Pro Trp Ala Pro His Pro Glu Tyr Gly Pro Pro Asn Asp Tyr
106 235                 240                 245                 250
108 ctg atc ctg aag ggc gag cac aag ttc aac tcg cag tac atc cgc gcg 1057
109 Leu Ile Leu Lys Gly Glu His Lys Phe Asn Ser Gln Tyr Ile Arg Ala
110                 255                 260                 265
112 gtg ggc ggc acg acg tgg cac tgg gcc gcg tcg gcg tgg cgc ttc att 1105
113 Val Gly Gly Thr Thr Trp His Trp Ala Ala Ser Ala Trp Arg Phe Ile
114                 270                 275                 280
116 ccg aac gac ttc aag atg aag agc gtg tac ggc gtc ggc cgc gac tgg 1153
117 Pro Asn Asp Phe Lys Met Lys Ser Val Tyr Gly Val Gly Arg Asp Trp
118             285                 290  .              295
120 ccg atc cag tac gac gat ctc gag ccg tac tat cag cgc gcg gag gaa 1201
121 Pro Ile Gln Tyr Asp Asp Leu Glu Pro Tyr Tyr Gln Arg Ala Glu Glu
122      300                 305                 310
124 gag ctc ggc gtg tgg ggc ccg ggc ccc gag gaa gat ctg tac tcg ccg 1249
125 Glu Leu Gly Val Trp Gly Pro Gly Pro Glu Glu Asp Leu Tyr Ser Pro
126 315                 320                 325                 330
128 cgc aag cag ccg tat ccg atg ccg ccg ctg ccg ttg tcg ttc aac gag 1297
```

RAW SEQUENCE LISTING   DATE: 05/19/2005
PATENT APPLICATION: US/10/511,796A   TIME: 10:15:53

Input Set : A:\Seqlist.txt
Output Set: N:\CRF4\05182005\J511796A.raw


```
129 Arg Lys Gln Pro Tyr Pro Met Pro Pro Leu Pro Leu Ser Phe Asn Glu
130             335                 340                 345
132 cag acc atc aag acg gcg ctg aac aac tac gat ccg aag ttc cat gtc   1345
133 Gln Thr Ile Lys Thr Ala Leu Asn Asn Tyr Asp Pro Lys Phe His Val
134             350                 355                 360
136 gtg acc gag ccg gtc gcg cgc aac agc cgc ccg tac gac ggc cgc ccg   1393
137 Val Thr Glu Pro Val Ala Arg Asn Ser Arg Pro Tyr Asp Gly Arg Pro
138             365                 370                 375
140 act tgt tgc ggc aac aac aac tgc atg ccg atc tgc ccg atc ggc gcg   1441
141 Thr Cys Cys Gly Asn Asn Asn Cys Met Pro Ile Cys Pro Ile Gly Ala
142     380                 385                 390
144 atg tac aac ggc atc gtg cac gtc gag aag gcc gaa cgc gcc ggc gcg   1489
145 Met Tyr Asn Gly Ile Val His Val Glu Lys Ala Glu Arg Ala Gly Ala
146 395                 400                 405                 410
148 aag ctg atc gag aac gcg gtc gtc tac aag ctc gag acg ggc ccg gac   1537
149 Lys Leu Ile Glu Asn Ala Val Val Tyr Lys Leu Glu Thr Gly Pro Asp
150             415                 420                 425
152 aag cgc atc gtc gcg gcg ctc tac aag gac aag acg ggc gcc gag cat   1585
153 Lys Arg Ile Val Ala Ala Leu Tyr Lys Asp Lys Thr Gly Ala Glu His
154             430                 435                 440
156 cgc gtc gaa ggc aag tat ttc gtg ctc gcc gcg aac ggc atc gag acg   1633
157 Arg Val Glu Gly Lys Tyr Phe Val Leu Ala Ala Asn Gly Ile Glu Thr
158             445                 450                 455
160 ccg aag atc ctg ctg atg tcc gcg aac cgc gat ttc ccg aac ggt gtc   1681
161 Pro Lys Ile Leu Leu Met Ser Ala Asn Arg Asp Phe Pro Asn Gly Val
162     460                 465                 470
164 gcg aac agc tcg gac atg gtc ggc cgc aac ctg atg gac cat ccg ggc   1729
165 Ala Asn Ser Ser Asp Met Val Gly Arg Asn Leu Met Asp His Pro Gly
166 475                 480                 485                 490
168 acc ggc gtg tcg ttc tat gcg agc gag aag ctg tgg ccg ggc cgc ggc   1777
169 Thr Gly Val Ser Phe Tyr Ala Ser Glu Lys Leu Trp Pro Gly Arg Gly
170             495                 500                 505
172 ccg cag gag atg acg tcg ctg atc ggt ttc cgc gac ggt ccg ttc cgc   1825
173 Pro Gln Glu Met Thr Ser Leu Ile Gly Phe Arg Asp Gly Pro Phe Arg
174             510                 515                 520
176 gcg acc gaa gcg gcg aag aag atc cac ctg tcg aac ctg tcg cgc atc   1873
177 Ala Thr Glu Ala Ala Lys Lys Ile His Leu Ser Asn Leu Ser Arg Ile
178             525                 530                 535
180 gac cag gag acg cag aag atc ttc aag gcc ggc aag ctg atg aag ccc   1921
181 Asp Gln Glu Thr Gln Lys Ile Phe Lys Ala Gly Lys Leu Met Lys Pro
182     540                 545                 550
184 gac gag ctc gac gcg cag atc cgc gac cgt tcc gca cgc tac gtg cag   1969
185 Asp Glu Leu Asp Ala Gln Ile Arg Asp Arg Ser Ala Arg Tyr Val Gln
186 555                 560                 565                 570
188 ttc gac tgc ttc cac gaa atc ctg ccg caa ccc gag aac cgc atc gtg   2017
189 Phe Asp Cys Phe His Glu Ile Leu Pro Gln Pro Glu Asn Arg Ile Val
190             575                 580                 585
192 ccg agc aag acg gcg acc gat gcg atc ggc att ccg cgc ccc gag atc   2065
193 Pro Ser Lys Thr Ala Thr Asp Ala Ile Gly Ile Pro Arg Pro Glu Ile
```

```
194           590                595                600  ·  ·
196 acg tat gcg atc gac gac tac gtg aag cgc ggc gcc gcg cat acg cgc   2113
197 Thr Tyr Ala Ile Asp Asp Tyr Val Lys Arg Gly Ala Ala His Thr Arg
198           605                610                615
200 gag gtc tac gcg acc gcc gcg aag gtg ctc ggc ggc acg gac gtc gtg   2161
201 Glu Val Tyr Ala Thr Ala Ala Lys Val Leu Gly Gly Thr Asp Val Val
202      620                625                630
204 ttc aac gac gaa ttc gcg ccg aac aat cac atc acg ggc tcg acg atc   2209
205 Phe Asn Asp Glu Phe Ala Pro Asn Asn His Ile Thr Gly Ser Thr Ile
206 635                640                645                650
208 atg ggc gcc gat gcg cgc gac tcc gtc gtc gac aag gac tgc cgc acg   2257
209 Met Gly Ala Asp Ala Arg Asp Ser Val Val Asp Lys Asp Cys Arg Thr
210           655                660                665
212 ttc gac cat ccg aac ctg ttc att tcg agc agc gcg acg atg ccg acc   2305
213 Phe Asp His Pro Asn Leu Phe Ile Ser Ser Ser Ala Thr Met Pro Thr
214           670                675                680
216 gtc ggt acc gta aac gtg acg ctg acg atc gcc gcg ctc gcg ctg cgg   2353
217 Val Gly Thr Val Asn Val Thr Leu Thr Ile Ala Ala Leu Ala Leu Arg
218           685                690                695
220 atg tcg gac acg ctg aag aag gaa gtc tgacc gtg cgg aaa tct act ctc  2403
221 Met Ser Asp Thr Leu Lys Lys Glu Val     Val Arg Lys Ser Thr Leu
222      700                705                710
224 act ttc ctc atc gcc ggc tgc ctc gcg ttg ccg ggc ttc gcg cgc gcg   2451
225 Thr Phe Leu Ile Ala Gly Cys Leu Ala Leu Pro Gly Phe Ala Arg Ala
226      715                720                725
228 gcc gat gcg gcc gat c                                              2467
229 Ala Asp Ala Ala Asp
230 730
233 <210> SEQ ID NO: 2
234 <211> LENGTH: 168
235 <212> TYPE: PRT
236 <213> ORGANISM: Burkhorderia cepacia
238 <400> SEQUENCE: 2
239 Met His Asn Asp Asn Thr Pro His Ser Arg Arg His Gly Asp Ala Ala
240 1               5                10               15
241 Ala Ser Gly Ile Thr Arg Arg Gln Trp Leu Gln Gly Ala Leu Ala Leu
242           20               25               30
243 Thr Ala Ala Gly Leu Thr Gly Ser Leu Thr Leu Arg Ala Leu Ala Asp
244      35               40               45
245 Asn Pro Gly Thr Ala Pro Leu Asp Thr Phe Met Thr Leu Ser Glu Ser
246      50               55               60
247 Leu Thr Gly Lys Lys Gly Leu Ser Arg Val Ile Gly Glu Arg Leu Leu
248 65               70               75               80
249 Gln Ala Leu Gln Lys Gly Ser Phe Lys Thr Ala Asp Ser Leu Pro Gln
250           85               90               95
251 Leu Ala Gly Ala Leu Ala Ser Gly Ser Leu Thr Pro Glu Gln Glu Ser
252           100                105                110
253 Leu Ala Leu Thr Ile Leu Glu Ala Trp Tyr Leu Gly Ile Val Asp Asn
254           115                120                125
```

```
255 Val Val Ile Thr Tyr Glu Glu Ala Leu Met Phe Gly Val Val Ser Asp
256     130                 135                 140
257 Thr Leu Val Ile Arg Ser Tyr Cys Pro Asn Lys Pro Gly Phe Trp Ala
258 145                 150                 155                 160
259 Asp Lys Pro Ile Glu Arg Gln Ala
260                 165
263 <210> SEQ ID NO: 3
264 <211> LENGTH: 539
265 <212> TYPE: PRT
266 <213> ORGANISM: Burkhorderia cepacia
268 <400> SEQUENCE: 3
269 Met Ala Asp Thr Asp Thr Gln Lys Ala Asp Val Val Val Val Gly Ser
270 1               5                 10                  15
271 Gly Val Ala Gly Ala Ile Val Ala His Gln Leu Ala Met Ala Gly Lys
272             20                  25                  30
273 Ala Val Ile Leu Leu Glu Ala Gly Pro Arg Met Pro Arg Trp Glu Ile
274             35                  40                  45
275 Val Glu Arg Phe Arg Asn Gln Pro Asp Lys Met Asp Phe Met Ala Pro
276     50                  55                  60
277 Tyr Pro Ser Ser Pro Trp Ala Pro His Pro Glu Tyr Gly Pro Pro Asn
278 65                  70                  75                  80
279 Asp Tyr Leu Ile Leu Lys Gly Glu His Lys Phe Asn Ser Gln Tyr Ile
280                 85                  90                  95
281 Arg Ala Val Gly Gly Thr Thr Trp His Trp Ala Ala Ser Ala Trp Arg
282             100                 105                 110
283 Phe Ile Pro Asn Asp Phe Lys Met Lys Ser Val Tyr Gly Val Gly Arg
284             115                 120                 125
285 Asp Trp Pro Ile Gln Tyr Asp Asp Leu Glu Pro Tyr Tyr Gln Arg Ala
286     130                 135                 140
287 Glu Glu Glu Leu Gly Val Trp Gly Pro Gly Pro Glu Glu Asp Leu Tyr
288 145                 150                 155                 160
289 Ser Pro Arg Lys Gln Pro Tyr Pro Met Pro Pro Leu Pro Leu Ser Phe
290                 165                 170                 175
291 Asn Glu Gln Thr Ile Lys Thr Ala Leu Asn Asn Tyr Asp Pro Lys Phe
292             180                 185                 190
293 His Val Val Thr Glu Pro Val Ala Arg Asn Ser Arg Pro Tyr Asp Gly
294             195                 200                 205
295 Arg Pro Thr Cys Cys Gly Asn Asn Asn Cys Met Pro Ile Cys Pro Ile
296     210                 215                 220
297 Gly Ala Met Tyr Asn Gly Ile Val His Val Glu Lys Ala Glu Arg Ala
298 225                 230                 235                 240
299 Gly Ala Lys Leu Ile Glu Asn Ala Val Val Tyr Lys Leu Glu Thr Gly
300                 245                 250                 255
301 Pro Asp Lys Arg Ile Val Ala Ala Leu Tyr Lys Asp Lys Thr Gly Ala
302             260                 265                 270
303 Glu His Arg Val Glu Gly Lys Tyr Phe Val Leu Ala Ala Asn Gly Ile
304             275                 280                 285
305 Glu Thr Pro Lys Ile Leu Leu Met Ser Ala Asn Arg Asp Phe Pro Asn
306     290                 295                 300
```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 05/19/2005
PATENT APPLICATION: US/10/511,796A TIME: 10:15:54

Input Set : **A:\Seqlist.txt**
Output Set: **N:\CRF4\05182005\J511796A.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:10; Xaa Pos. 6,17,18,19,22
Seq#:17; Xaa Pos. 2,3

VERIFICATION SUMMARY   DATE: 05/19/2005
PATENT APPLICATION: US/10/511,796A   TIME: 10:15:54

Input Set : **A:\Seqlist.txt**
Output Set: **N:\CRF4\05182005\J511796A.raw**

```
L:36 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:1
L:39 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:1
L:413 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:417 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:10
L:421 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:10
L:425 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:10
L:426 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:0
M:341 Repeated in SeqNo=10
L:846 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:850 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:17
L:851 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17 after pos.:0
```